

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-250575

(43)Date of publication of application : 14.09.2000

(51)Int.Cl.

G10L 15/00
G10L 15/18
H04N 5/44
H04N 7/025
H04N 7/03
H04N 7/035
H04N 7/173

(21)Application number : 2000-055787

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 01.03.2000

(72)Inventor : JUNQUA JEAN-CLAUDE
CONTOLINI MATTEO

(30)Priority

Priority number : 99 260156

Priority date : 01.03.1999

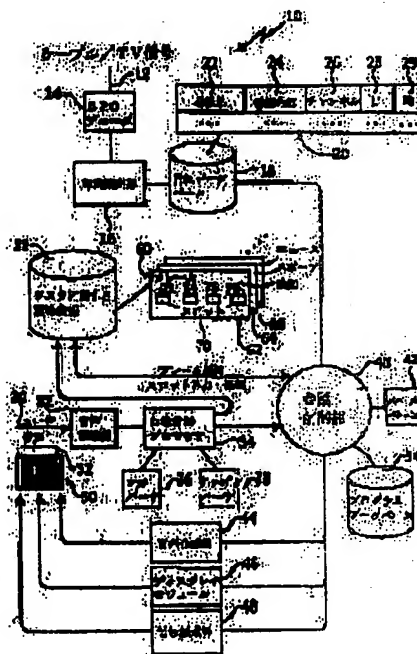
Priority country : US

(54) SPEECH UNDERSTANDING DEVICE AND METHOD FOR AUTOMATICALLY SELECTING BIDIRECTIONAL TV RECEIVER

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a speech understanding system capable of automatically selecting a television program by receiving a request dictated by a user and processing the request for a knowledge base related to program information.

SOLUTION: This speech understanding system is provided with a knowledge extraction part 16 for receiving/processing EPG(electronic programing guide) information and forming a program database, a speech recognition part 32 for translating the dictated request to a series of text information, a natural language processor 34 for receiving a series of text information and explaining the meaning of a word contents of the dictated request and a conversation control part 40 for judging that a key word slot is inputted sufficiently and further asking for the user when a free slot exists. Thus, the program database is retrieved by using the key word arranged in a task frame, and the program is selected, and a television channel related to the television program is selected.



【特許請求の範囲】

【請求項1】 ユーザが口述するリクエストを受け、このリクエストをプログラム情報に関する知識ベースに対して処理して、テレビ番組を自動的に選局する音声理解システムであって、

電子番組ガイド（EPG）の情報を受信し、EPG情報を処理して、番組データベースを形成する知識抽出部と、

口述リクエストを受け、口述リクエストを複数の単語からなる一連のテキスト情報に翻訳する音声認識部と、
一連のテキスト情報を受け、口述リクエストの語義内容を解釈するように単語を処理する自然言語プロセッサと、

十分な数のキーワードスロットが入力されたかどうかを判断するためにタスクフレームを分析し、空のスロットに入力するためにユーザに対して追加的情報を質問するための会話制御部であって、

タスクフレーム内に配置されたキーワードを用いて番組データベースを検索して番組を選択し、テレビ番組に関連するテレビチャンネルを選局するための信号を形成する会話制御部と、を備えることを特徴とする音声理解システム。

【請求項2】 自然言語プロセッサは、一連のテキスト情報内に含まれる単語を分析し、キーワードを特定するためのローカルパーサを有し、このローカルパーサが、LR文法データベースを用いて、単語の意味を解釈することを特徴とする請求項1の音声理解システム。

【請求項3】 ローカルパーサが、各キーワードに対してタグ付きデータ構造を形成し、このタグ付きデータ構造がキーワードの意味を表現することを特徴とする請求項2の音声理解システム。

【請求項4】 自然言語プロセッサは、各キーワードに対するタグ付きデータ構造を受けるとともに、口述リクエストの意味に付随するタスクフレームを選択するためのグローバルパーサを有することを特徴とする請求項3の音声理解システム。

【請求項5】 グローバルパーサが、口述リクエストの意味に付随するタスクフレームを決定するために複数の決定樹木と相互通信することを特徴とする請求項4の音声理解システム。

【請求項6】 グローバルパーサが、各キーワードに対するタグ付きデータ構造を受け、キーワードスロットにキーワードの意味を配置することを特徴とする請求項4の音声理解システム。

【請求項7】 会話制御部は、タスクフレームが番組データベースを検索する上で十分な情報を有していることを判断するルールベースと、相互通信することを特徴とする請求項1の音声理解システム。

【請求項8】 会話制御部は、追加的情報を得るために、音声合成部を介してユーザに質問することを特徴と

する請求項1の音声理解システム。

【請求項9】 会話制御部は、追加的情報を得るために、表示システムを介してユーザに質問することを特徴とする請求項1の音声理解システム。

【請求項10】 会話制御部から指令を受け、リモート受信器を操作するための信号を形成するための信号形成部を有することを特徴とする請求項1の音声理解システム。

【請求項11】 番組データベースは、ユーザが選択できる番組とチャンネルの組み合わせに関する複数の番組レコードを有することを特徴とする請求項1の音声理解システム。

【請求項12】 ユーザが口述するリクエストを受け、このリクエストをプログラム情報に関する知識ベースに対して処理して、テレビ番組を自動的に選局する音声理解システムであって、

電子番組ガイド（EPG）の情報を受信し、EPG情報を処理して、番組データベースを形成する知識抽出部と、

口述リクエストを受け、口述リクエストを複数の単語からなる一連のテキスト情報に翻訳する音声認識部と、
一連のテキスト情報を受け、複数のキーワードからなるタスクフレームにより表現される口述リクエストの語義内容を解釈するように単語を処理する自然言語プロセッサであって、

文法構造に関するデータベースを用いて単語を分析し、単語の意味を表現するタグ付きデータ構造を形成するためのローカルパーサを有し、

タスクフレームを選択し、タグ付きデータ構造に含まれる単語の意味をタスクフレーム内のキーワードスロットに配置するためのグローバルパーサを有する自然言語プロセッサと、

十分な数のキーワードスロットが入力されたかどうかを判断するためにタスクフレームを分析し、所定の組み合わせのスロットが入力されるまでユーザに追加的情報を質問するための会話制御部であって、

タスクフレーム内に配置されたキーワードを用いて番組データベースを検索して番組を選択し、テレビ番組に関連するテレビチャンネルを選局するための信号を形成する会話制御部と、を備えることを特徴とする音声理解システム。

【請求項13】 グローバルパーサが、各キーワードに対するタグ付きデータ構造を受け、口述リクエストの意味に付随するタスクフレームを選択することを特徴とする請求項12の音声理解システム。

【請求項14】 グローバルパーサが、口述リクエストの意味に付随するタスクフレームを決定するために複数の決定樹木と相互通信することを特徴とする請求項12の音声理解システム。

【請求項15】 各タグ付けされた単語に対する各タグ

付きデータ構造が、どのタスクフレームが単語に付随するかを示すフレームタグと、どのキーワードスロットが単語の意味に付随するかを示すスロットタグと、を有することを特徴とする請求項12の音声理解システム。

【請求項16】 会話制御部は、タスクフレームが所定の組み合わせの入力スロットを有していることを判断するルールベースと、相互通信することを特徴とする請求項12の音声理解システム。

【請求項17】 テレビ信号、ケーブル信号、衛星放送テレビ信号および電話回線による信号のいずれか1つからEPG情報を受信することを特徴とする請求項12の音声理解システム。

【請求項18】 口述リクエストを理解し、口述リクエストに基づいてテレビチャンネルを選局する方法であって、

電子番組ガイド(EPG)の情報を受信するステップと、

番組データベースを形成するためにEPG情報を処理するステップと、

ユーザから口述リクエストを受け、ユーザが要求するタスクの意味表現集を抽出するために、自然言語プロセッサを用いて口述リクエストを処理するステップと、

自然言語プロセッサにより特定されたキーワードの検索リクエストを構築するステップと、

キーワードを用いて番組データベースを検索するステップと、

所望する番組に関連するテレビチャンネルを選局するための信号を形成するステップと、を有することを特徴とする方法。

【請求項19】 さらに、ユーザに対して追加的に質問するために、会話制御部を用いて対話処理を行うステップを有することを特徴とする請求項18の方法。

【請求項20】 さらに、文法構造に関するデータベースを用いてユーザのリクエストを分析し、口述リクエストの意味を表現する1つまたはそれ以上のタグ付きデータ構造を形成するステップを有することを特徴とする請求項19の方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、双方向性テレビ受像機において、チャンネル選局処理を自動的に行うための音声理解装置および音声理解方法に関する。とりわけ、本発明は、口頭によるチャンネル選局のリクエスト(要求)を受け、チャンネルを自動的に選局するために番組検索放送(電子番組ガイド: EPG, Electronic Program Guide)から抽出された情報に対してこのリクエストを処理するための自然言語処理技術を用いたシステムに関する。

【0002】

【発明の背景】双方向性テレビ受像機は、双方向のコミ

ュニケーションが可能となるように設計されている。つまり、従来式の一方向性テレビ受像機(以下、単に「TV」という。)やラジオ放送受信機とは異なり、通常の双方向性TVを用いるとき、視聴者は電話回線を介して応答することができる。また、デジタル式テレビ放送によれば、現在のところ1つのアナログチャンネルしか送信できない空間にいくつかのテレビ放送チャンネルを送信することかでき、これは新規でより効率的なテレビ放送の手法である。これにより、より多くの新しいチャンネルやテレビ番組を提供する機会が与えられる。

【0003】将来において、すべてのテレビ放送がデジタル式に放送されることが予想される。デジタル技術を用いることにより、およそ200もの、あるいはより以上のチャンネルを利用できるようになる。デジタル放送により、この他にも例えば、高品位またはワイドスクリーンの画像、CD並に良質な音声、異なるチャンネルで別々の時間に始まる映画を視聴者が都合のよい時に鑑賞できる「ビデオ・オン・デマンド」など、視聴者に多くの利便が得られることが期待される。例えば、ホームバンキング、ホームショッピング、およびインターネットへの接続といった双方向性サービスも同様に、テレビ受像機を介して、デジタル式に提供されるようになるだろう。

【0004】何らかのデジタル技術を用いて、視聴者がデジタル放送を受信するように選局するとき、特別のデジタル受信機、またはテレビ用受信端末(据置型筐体: set-top box)のデコーダを購入するか賃借する必要がある。このデコーダによりデジタル放送画像をスクリーン上に再形成することができる。このデコーダが組み込まれたテレビ受像機も市販されるであろう。テレビ用受信端末には、音声認識および音声理解などの追加的サービスを提供するために用いることができるプロセッサが設けられている。

【0005】テレビチャンネルの数が増えれば増えるほど、視聴者にとってチャンネル選局が難しくなる。本発明によれば、リストにあるチャンネルすべてを1つずつ選局して所望のチャンネルを探すのではなく、口述されたリクエストの内容を分析して、視聴者が見たい番組の諸義内容を理解することにより、チャンネルを選局することができる。視聴者は、自然な言語および会話を用いてリクエストを声に出し、所望する番組およびテレビ放送チャンネルを選局することができる。

【0006】上述のように、数多くのテレビ信号の1つとして放送される、番組検索放送(電子番組ガイド: EPG, Electronic Programming Guide, 以下、単に「EPG」という。)の情報に関する知識表現(Knowledge Representation)を形成し、この情報をTVまたはテレビ用受信端末に記憶させることが好ましい。ユーザの口頭による番組選局リクエストの意味を抽出し、EPGの知識データベースに対してこのリクエストを照合し、所

望の番組を検索し、テレビを対応するチャンネルに選局できるシステムも同様に備えておくことが好ましい。最後に、ユーザのリクエストを明確にし、曖昧さを取り除くための対話処理（ダイアログフェーズ：dialog phase）を行い、ユーザと情報交換するシステムも同様に備えておくことが好ましい。

【0007】（発明の要約）本発明の方法および装置は、放送内容および番組検索放送（EPG）を検索した後、テレビ受像機またはビデオ録画装置のチャンネルを自動的に選局するために音声理解技術を用いる。この技術によれば、自然な言語および会話による口述リクエストに基づいて、希望するTV番組を選局する。映画などのいくつかの内容の要約とともにTV番組情報を含むEPGの内容を用いて、動的な認識語彙集が形成される。この語彙集は、システムが認識し理解できるキーワードとフレーズを有している。この語彙集は、その他に、ユーザがチャンネルを選局するときによく用いられる非キーワード（例えば、「I」、「would」または「please」）を有している。

【0008】自然言語プロセッサは、口述リクエストの語義内容を理解するために、例えば「スポーツ（sport）」および「映画（movie）」といったキーワードの意味に関する仮説知識（a priori knowledge）を含む文法データ構造からなるデータベースを用いる。さらに、会話制御部は、チャンネル選局タスクの意味に関する表現集（Representation）を用いて、ユーザにリクエストを完成させるよう支援する。自然言語プロセッサは、ローカル構文解釈モジュールおよびグローバル構文解釈モジュールを用いて、口述リクエストの意味論上の関連する部分を特定し、会話制御部にこの特定情報を送る。また、この自然言語プロセッサは、一旦理解された単語およびフレーズを、口述リクエストの意味表現集として組織化する。

【0009】チャンネル選局の意味表現集は、例えば、選局すべき番組、時間、おそらくチャンネルまたはネットワークなどの情報を備えている。視聴者は、チャンネルを選局することができるが、同様に、特定の日に視聴できる番組に関する情報を質問することができる。この情報は、音声を用いて応答するか、またはテレビ画面を用いて表示することができる。自然言語プロセッサを用いるので、視聴者は、音声理解装置と対話する際に用いる言葉のスタイルを気遣うことなく、自らの目的を達成することに専念することができる。曖昧なところがあるとき、ユーザがさらにリクエストを洗練できるように、分割画面モードを用いるようにしてもよい。

【0010】ユーザの好みに関する履歴データベースが、会話制御部により自動的に構築される。履歴データベースは、過去の口述リクエストまたは会話を含んでもよい。この過去の口述リクエストまたは会話を用いて、会話制御部は、認識処理中にいくつかの単語を優先

（favor）するように、音声理解装置の言語モデルを動的に変更することができる。この履歴データベースは、同様に、対話処理中に、ユーザの好みに関して質問するときの質問のランク付けを行うために用いることができる。

【0011】この音声理解装置を用いて、チャンネル選局をビデオ録画装置の録画コマンドに接続することにより、ビデオ録画装置をプログラムすることができる。頻繁に録画する番組のために、ユーザに対する追加的機能としてマクロ命令を確立するようにしてもよい。同様に、例えば、「バスケットの試合は明日放送される」などと、視聴者に対してリマインドを出して周知させることができる。

【0012】本発明の追加的な対象物、利点、および特徴が、以下の明細書および添付されたクレームに添付する図面を参照すると明らかとなる。

【0013】

【課題を解決するための手段】本発明の請求項1に係る音声理解システムは、ユーザが口述するリクエストを受け、このリクエストをプログラム情報に関する知識ベースに対して処理して、テレビ番組を自動的に選局する音声理解システムであって、電子番組ガイド（EPG）の情報を受信し、EPG情報を処理して、番組データベースを形成する知識抽出部と、口述リクエストを受け、口述リクエストを複数の単語からなる一連のテキスト情報に翻訳する音声認識部と、一連のテキスト情報を受け、口述リクエストの語義内容を解釈するように単語を処理する自然言語プロセッサと、十分な数のキーワードスロットが入力されたかどうかを判断するためにタスクフレームを分析し、空のスロットに入力するためにユーザに対して追加的情報を質問するための会話制御部であって、タスクフレーム内に配置されたキーワードを用いて番組データベースを検索して番組を選択し、テレビ番組に関連するテレビチャンネルを選局するための信号を形成する会話制御部と、を備える。

【0014】本発明の請求項2に係る音声理解システムにおいて、自然言語プロセッサは、一連のテキスト情報内に含まれる単語を分析し、キーワードを特定するためのローカルパーサを有し、このローカルパーサが、LR文法データベースを用いて、単語の意味を解釈する。

【0015】本発明の請求項3に係る音声理解システムにおいて、ローカルパーサが、各キーワードに対してタグ付きデータ構造を形成し、このタグ付きデータ構造がキーワードの意味を表現する。

【0016】本発明の請求項4に係る音声理解システムにおいて、自然言語プロセッサは、各キーワードに対するタグ付きデータ構造を受けるとともに、口述リクエストの意味に付随するタスクフレームを選択するためのグローバルパーサを有する。

【0017】本発明の請求項5に係る音声理解システム

において、グローバルバーサが、口述リクエストの意味に付随するタスクフレームを決定するために複数の決定樹木と相互通信する。

【0018】本発明の請求項6に係る音声理解システムにおいて、グローバルバーサが、各キーワードに対するタグ付きデータ構造を受け、キーワードスロットにキーワードの意味を配置する。

【0019】本発明の請求項7に係る音声理解システムにおいて、会話制御部は、タスクフレームが番組データベースを検索する上で十分な情報を有していることを判断するルールベースと、相互通信する。

【0020】本発明の請求項8に係る音声理解システムにおいて、会話制御部は、追加的情報を得るために、音声合成部を介してユーザに質問する。

【0021】本発明の請求項9に係る音声理解システムにおいて、会話制御部は、追加的情報を得るために、表示システムを介してユーザに質問する。

【0022】本発明の請求項10に係る音声理解システムにおいて、会話制御部から指令を受け、リモート受信器を操作するための信号を形成するための信号形成部を有する。

【0023】本発明の請求項11に係る音声理解システムにおいて、番組データベースは、ユーザが選択できる番組とチャンネルの組み合わせに関する複数の番組レコードを有する。

【0024】本発明の請求項12に係る音声理解システムは、ユーザが口述するリクエストを受け、このリクエストをプログラム情報に関する知識ベースに対して処理して、テレビ番組を自動的に選局する音声理解システムであって、電子番組ガイド(EPG)の情報を受信し、EPG情報を処理して、番組データベースを形成する知識抽出部と、口述リクエストを受け、口述リクエストを複数の単語からなる一連のテキスト情報に翻訳する音声認識部と、一連のテキスト情報を受け、複数のキーワードからなるタスクフレームにより表現される口述リクエストの語義内容を解釈するように単語を処理する自然言語プロセッサであって、文法構造に関するデータベースを用いて単語を分析し、単語の意味を表現するタグ付きデータ構造を形成するためのローカルバーサを有し、タスクフレームを選択し、タグ付きデータ構造に含まれる単語の意味をタスクフレーム内のキーワードスロットに配置するためのグローバルバーサを有する自然言語プロセッサと、十分な数のキーワードスロットが入力されたかどうかを判断するためにタスクフレームを分析し、所定の組み合わせのスロットが入力されるまでユーザに追加的情報を質問するための会話制御部であって、タスクフレーム内に配置されたキーワードを用いて番組データベースを検索して番組を選択し、テレビ番組に関連するテレビチャンネルを選局するための信号を形成する会話制御部と、を備える。

【0025】本発明の請求項13に係る音声理解システムにおいて、グローバルバーサが、各キーワードに対するタグ付きデータ構造を受け、口述リクエストの意味に付随するタスクフレームを選択する。

【0026】本発明の請求項14に係る音声理解システムにおいて、グローバルバーサが、口述リクエストの意味に付随するタスクフレームを決定するために複数の決定樹木と相互通信する。

【0027】本発明の請求項15に係る音声理解システムにおいて、各タグ付けされた単語に対する各タグ付きデータ構造が、どのタスクフレームが単語に付随するかを示すフレームタグと、どのキーワードスロットが単語の意味に付随するかを示すスロットタグと、を有する。

【0028】本発明の請求項16に係る音声理解システムにおいて、会話制御部は、タスクフレームが所定の組み合わせの入力スロットを有していることを判断するルールベースと、相互通信する。

【0029】本発明の請求項17に係る音声理解システムにおいて、テレビ信号、ケーブル信号、衛星放送テレビ信号および電話回線による信号のいずれか1つからEPG情報を受信する。

【0030】本発明の請求項18に係るテレビチャンネルを自動的に選局する方法は、口述リクエストを理解し、口述リクエストに基づいてテレビチャンネルを選局する方法であって、電子番組ガイド(EPG)の情報を受信するステップと、番組データベースを形成するためにEPG情報を処理するステップと、ユーザから口述リクエストを受け、ユーザが要求するタスクの意味表現集を抽出するために、自然言語プロセッサを用いて口述リクエストを処理するステップと、自然言語プロセッサにより特定されたキーワードの検索リクエストを構築するステップと、キーワードを用いて番組データベースを検索するステップと、所望する番組に関連するテレビチャンネルを選局するための信号を形成するステップと、を有する。

【0031】本発明の請求項19に係るテレビチャンネルを自動的に選局する方法において、さらに、ユーザに対して追加的に質問するために、会話制御部を用いて対話処理を行うステップを有する。

【0032】本発明の請求項20に係るテレビチャンネルを自動的に選局する方法において、さらに、文法構造に関するデータベースを用いてユーザのリクエストを分析し、口述リクエストの意味を表現する1つまたはそれ以上のタグ付きデータ構造を形成するステップを有する。

【0033】

【発明の実施の形態】本発明の技術によれば、音声理解技術を用いて、テレビチャンネルを自動的に選局するシステムおよび方法が開示されている。図1は、本発明の好適な実施例による音声理解・チャンネル選局システム

を示す。チャンネル選局システム10は、テレビ用受信端末デコーダ52に内蔵されていることが好ましい。しかし、このシステム10は、テレビ50や衛星放送チューナ、あるいはビデオ再録装置に内蔵することもできる。

【0034】本発明の一部として、ケーブル放送またはテレビ放送信号12は、チャンネル選局システム10に対する番組検索放送（電子番組ガイド：EPG, Electronic Programming Guide）の情報を含む。またEPG情報は、電話通信回線を介してインターネット系サービスプロバイダ、または専用のダイヤル呼出式EPGサービスプロバイダからダウンロードすることもできる。テレビ信号12は、再生して見ることもできるし、そして／または記録することもできる。EPGデコーダ14は、EPG情報を受信して、必要ならば表示する。EPGデコーダ14は、EPG情報をテキスト情報に変換およびフォーマットし、このテキスト情報が、知識抽出部18に通信される。知識抽出部18は、このEPG情報を番組データベース18内に格納し検索できるフォーマットに再構成する機能を有する。

【0035】この番組データベース18は、予め決められた一連の項目に関する検索可能な番組レコードを数多く保有していて、この番組レコードは、例えば、これに限るわけではないが、番組名の項目22、番組の説明または主題の項目24、チャンネルの項目26、日付の項目28、時間項目29などに関する。番組データベース18の番組レコード20は、EPGの情報内容が変わると、新しいものに随時更新される。したがって、番組データベース18が更新されるのを待たずに、いつでも、口述リクエストを処理することができる。加えて、番組データベース18に記録された、時間経過して意味のなくなった番組レコード20は、周期的に取り除かれるので、限られた制御できる数の番組レコード20だけを、チャンネル選局システム10を用いて、ユーザの口述リクエストを満足するように検索すればよい。

【0036】一方、チャンネル選局システム10の音声処理に関し、符号30でユーザの音声として示した口述リクエストおよび口述情報が、音声認識部32に入力される。口述された単語は、音声認識部32により処理され、テキスト情報に変換される。適当な音声認識装置が、リー・Kによる「語彙語者不特定連続音声認識（The Sphinx Systems, Ph.D. Thesis, Carnegie Mellon University, 1988）」に教示されている。音声認識部32から出力された一連のテキスト情報は自然言語プロセッサ34に供給され、この自然言語プロセッサ34は、主に、一連のテキスト情報を分析して、語義内容および口述リクエストの意図を理解する機能を有する。自然言語プロセッサ34による音声理解分析は、ローカルパーサ（ローカル構文解釈プログラム）モジュール36とグローバルパーサ（グローバル構文解釈プログラム）モジュ

ール38により実行される。自然言語プロセッサ34とその構成部品については、後により詳しく説明する。

【0037】演算処理装置を用いた会話制御部40は、自然言語プロセッサ34を含むチャンネル選局システム10の種々のモジュールと相互通信する。会話制御部40は、要求のあったタスク（課題）に関する意味表現集を有するタスクデータベース58と情報交換する。タスクデータベース58は、あらかじめ決められた複数のタスクフレーム60を有し、このタスクフレーム60は、ユーザの口述リクエストに関連したタスク（課題）の意味表現集を有する。図示するように、このタスクフレーム60は、映画タスクフレーム62、スポーツタスクフレーム64、ニュースタスクフレーム66を有する。ここでは、タスクフレーム60を3つしか図示しないが、本発明においては、その他数多くのタスクフレームを設けることができると理解すべきである。各タスクフレーム60は、ユーザの口述リクエストから構文解釈されたキーワードの理解された意味を記憶するために、複数のキーワードスロット70を備えている。後により詳しく説明するが、自然言語プロセッサ34内に設けたグローバルパーサモジュール38により、これらのフレームおよびスロットにデータが格納される。

【0038】会話制御部40は、選択されたタスクフレーム60に含まれる検索基準を用いて、番組データベース18から番組レコード20を呼び出す。会話制御部40が実行する検索機能は、後により詳しく説明するルールベース42により支援される。リクエスト履歴データベース54は、例えば、鑑賞および／または記録したい好みのスポーツや映画のタイプなど、ユーザの好みに関する履歴を記憶するために、会話制御部40により維持管理される。

【0039】会話制御部40は、音声合成部44に出力することができ、この音声合成部44は、ユーザに対して音声で質問することができる。会話制御部40は、画面上で設定可能なディスプレイ（オンスクリーンディスプレイ：OSD, On Screen Display）モジュール46に出力して、接続されたテレビ画面50を介してユーザに対して質問を表示することができる。さらに、会話制御部40は、信号形成部48に出力信号を送ることができ、信号形成部48は、この出力信号を、テレビ受像機50またはテレビ用受信端末52をチャンネル変更させるような適当な信号に変換することができる。本発明の一部をなすと考えるべきことであるが、信号形成部48は、通常のテレビ受像機、ケーブル受信端末、衛星放送受信機、ビデオ録画装置で用いられるリモート指令受信器と互換性のある、広く用いられた赤外線信号を形成することができる。このようにして、会話制御部40は、自動的にテレビチャンネルを変更したり、特定の日時に所望のチャンネルから番組を録画するように、ビデオ録画装置をプログラムすることもできる。

【0040】図2は、自然言語プロセッサ34の動作を示す。上述のとおり、自然言語プロセッサ34は、音声認識部32により提供されるデジタル化した単語の語義内容をさらに分析して理解するために、ローカルバーサ36とグローバルバーサ38を備える。ローカルバーサ36は、単語、フレーズ、段落文章、およびその他すべてのタイプの口述された文法的表現を分析する機能を有する。自然言語プロセッサ34を簡単に説明するために、以下、認識され理解される文法的表現とは、すべて単語に関するものとする。したがって、単語に関して説明されたことは、フレーズ、段落文章、その他すべてのタイプの口述された文法的表現に関するものと理解されたい。

【0041】ローカルバーサ36は、LR文法モジュール86を用いて単語を吟味し、その単語がキーワードかそうでないかを判断する。単語がキーワードであると認識されたとき、その単語（またはフレーズなど）に、単語の理解される意味を表現するデータ構造を「タグ付け」する。こういった吟味は、システム語彙集で構成される文法データ構造のデータベースを用いてなされる。つまり、理解可能な単語またはフレーズには、その単語に対するタグを表現する、文法データ構造が付随する。ローカルバーサ36を用いて、正しい文法データ構造を特定すると、その単語に対して、例えば、タグ付けデータ構造82または84などのタグ付けデータ構造を形成して、その単語の意味を定義する。ローカルバーサ36は、口述され、キーワードであると特定された単語のすべてについて、適当なタグ付けデータ構造を付与する機能を有する。グローバルバーサ38は、選択されたタスクフレーム60のキーワードスロット70にタグ付けされた単語すべてを配置する機能を有する。

【0042】動作について説明すると、ローカルバーサ36が各単語を受け取り、LR文法モジュール86を用いて、その単語に付随する文法データ構造を検索する。その単語に対する文法データ構造により、ローカルバーサ36は、その単語がキーワードかどうかを判断し、適当なタグ付けデータ構造82、84をどのように形成すべきか了解する。その単語がキーワードでなかった場合、グローバルバーサ38によりさらに分析する必要がある場合に備えて、この単語はバッファ内に格納される。その単語がキーワードであった場合、その文法データ構造が、タグ付けデータ構造の形成方法に関する情報を含んでいる。その単語がキーワードでなかった場合、フレームタグおよびスロットタグの項目は空の状態であり、キーワードでない単語がバッファに格納される。グローバルバーサ38は、このフレームおよびスロットに関するタグ情報を用いて、適当なタスクフレーム60の適当なスロット70にキーワードを配置することができる。この処理は、フレーム選択・スロット入力モジュール56により支援される。あるキーワードを複数のフレ

ームに配置できる場合、2つの異なるタスクフレーム60の同じスロット70に同じキーワードを入力する必要があることを、タグ付けデータ構造82、84が指示する。正しいタスクフレーム60は、グローバルバーサ38による後の反復処理の間に選択することができる。

【0043】口述されたリクエストが、例えば、「私は、今夜、映画を見たい。(I would like to watch a movie.)」であったとする。この例示したリクエストは、いくつかのキーワード、すなわち、「見る (watch)」、「映画 (movie)」、および「今夜 (tonight)」を有している。残りの単語は、キーワードでないと仮定する。しかし、ユーザはどんな映画を見たいのか、そしてユーザは何時から鑑賞（記録）し始めたいのかについて、具体的に知るためには、この例示的なリクエストに関して対話処理（ダイアログフェーズ）を行う必要がある。この分析技術の一部であるが、ローカルバーサ36は、単語「私 (I)」、「would」、「like」、および「to」を個別に処理し、これらの単語がキーワードでないことを確認すると、これらの非キーワードの単語をバッファ（図示せず）に格納する。そこで、ローカルバーサ36が、単語「鑑賞する (watch)」に関する文法データ構造を呼び出して、タグ付きデータ構造を形成し、この単語にタグ付きデータ構造を付与する。タグが付与された単語は、その後、グローバルバーサ38に伝えられ、このグローバルバーサ38は、ユーザの希望する動作が、番組を録画することではなくて、鑑賞することであることを確認し、あるいは、鑑賞したい将来の日時について質問する。

【0044】単語「映画 (movie)」に関するタグ付きデータ構造（データ構造82として示す）は、映画タスクフレーム62を選択する必要があることを指示する。しかし、このキーワードはタスクフレーム60に相当強く付随しているので、キーワードスロット70に「映画 (movie)」が指定されない。単語「今夜 (tonight)」に関するタグ付きデータ構造は、このキーワードに関する意味表現集が、すべてのタスクフレーム60の日付スロット78内に配置されなければならないことを指示する。しかし、グローバルバーサ38は、映画タスクフレーム62の日付スロット78に、この単語「今夜」の理解される意味を入力する。さらに、タグ付きデータ構造は、映画タスクフレーム62の時間スロット80に、午後5:00以降の時間を入力する必要があることを指示するようにしてもよい。こうして、会話制御部40は、ユーザが今日の日の夕方に番組をサ検索したいことを認識する。

【0045】この時点で、ローカルバーサ36は、口述リクエスト内のすべての単語についてタグ付けを行い、グローバルバーサ38は、フレーム選択・スロット入力モジュール56と協働して、サーチリクエストを構築するために適当なタスクフレーム60を選択し、適当なス

10

20

30

40

50

ロット70に単語の理解される意味を入力した。次に、会話制御部40は、より具体的な情報について質問することができる。会話制御部40は、映画タスクフレーム62内の入力しなければならないキーワードスロット70から、ユーザに対して質問すべきことを了解している。例えば、時間スロット80が空の状態ならば、会話制御部40は、「何時から、あなたは映画を見たいのですか。(at what time would you like to watch a movie?)」と、ユーザに尋ねる。ユーザが口頭で時刻や時間範囲を応えた場合、ローカルバーサ38は、上述の技術を用いて時間に関するキーワードをタグ付けし、グローバルバーサ38は、これらのキーワードを映画タスクフレーム62の時間スロット80内に配置する。

【0046】グローバルバーサ38は、主に、ローカルバーサ36により形成されたタグ付きデータ構造を分析して、口述されたリクエストの文脈におけるその単語の意味を特定し、適当なスロット70に単語の意味を配置する機能を持つ。グローバルバーサ38は、数多くの決定樹木構造88からなる。口述された要請の文脈が決定されると、特定の決定樹木88が用いられる。決定樹木88は、開始点を有し、特定の処理をもって終了する。決定樹木88の終了点における処理は、単語をどこに配置すべきか、あるいは特定の曖昧さをどのように解決すべきかについて、グローバルバーサに指示することである。本発明の場合、この終了点における処理とは、一般に、どのタスクフレーム60を選択すべきか、どのキーワードスロット70に特定のタグ付き単語を配置すべきかについて、グローバルバーサ38に指示することである。

【0047】入力されたキーワードスロット70のどの組み合わせを用いれば、番組データベース18内をサーチするための十分な情報が提供されるかについて、会話制御部40が決定する際に、ルールベース42がこれを支援する。例えば、映画タスクフレーム62の時間キーワードスロット80が入力されると、会話制御部40は、リクエストのあった時間または時間帯に始まるすべての映画を検索することができる。しかし、検索の結果、リクエストのあった時間に予定した数以上の映画が確認された場合、会話制御部40は、「どんなタイプの映画を見たいですか。(what type of movie would you like to watch?)」と、ユーザに尋ねるようにしてもよい。このとき、会話制御部40は、映画タスクフレーム62内の主題キーワードスロット74またはジャンルキーワードスロット78に入力しようとしている。ユーザが、口頭により、主題またはジャンルを応えた場合、ローカルバーサ38は、上述の技術を用いて、主題またはジャンルに関するキーワードをタグ付けする。これらの新たにタグ付けされた単語は、その後、グローバルバーサ38に伝えられ、映画タスクフレーム62内の適当なスロット70に配置される。

【0048】例えば、ユーザが、「アクション映画(action movie)」と応えた場合、グローバルバーサ38は、単語「アクション(action)」をジャンルスロット78に配置し、会話制御部40は、再び絞り込み検索を行う。希望する時間にアクション映画が1つしかなかった場合、会話制御部40は、テレビ受信機50またはテレビ用受信端末52を番組レコード20内のチャンネルに自動的に切り換えるように信号形成部48に指示する。希望する時間にいくつかのアクション映画が鑑賞できる場合、会話制御部40は、オンスクリーンディスプレイモジュール46を介して、鑑賞可能なすべての映画を列挙するようにしてもよい。このとき、ユーザは、番号かチャンネルを用いて希望する映画を選択するようにしてもよい。本発明の別の特徴として、会話制御部40は、チャンネルを切り換える前にユーザに対してユーザリクエストをフィードバックして、確認を取るようにしてもよい。

【0049】例えば、「私は、今夜、デトロイト・レッド・ウィングズのホッケーの試合を見たい(I would like to watch the Detroit Red Wings hockey game tonight.)」とか、「私は、今週の火曜日に、PBSで放送されるノバという番組を録画したい(I would like to record the program Nova on PBS this Tuesday.)」などと、1つの口述リクエストによって十分な一連の情報を会話制御部40に提供することにより、ユーザが慣れてくると、自然言語プロセッサ34により、十分なキーワードスロット70を入力して、検索を実行することができる。会話制御部40にとって、口述リクエストは十分満足できるものとなる。リクエストがホッケーの試合を鑑賞することである場合、会話制御部40は、検索を完了して、番組レコード20のチャンネル項目26内に含まれる情報に基づいて、テレビ用受信端末52またはテレビ受信機50のチャンネルを変えるための適当な信号を形成することができる。リクエストが、PBSから希望する番組を記録することである場合、会話制御部40は、検索を完了して、番組レコード20から日付、時間、およびチャンネルの情報を呼び出して、信号形成部を介して適当なビデオ録画装置をプログラムするための信号を形成する。

【0050】ビデオ録画装置が同じ時刻に別の番組を録画するようにすでにプログラムされている場合、またはビデオ録画装置に空のテープを挿入する必要がある場合、本発明の一部として、会話制御部は、ビデオ録画装置からフィードバック信号を受信するように構成される。こうして、ユーザがいる場合に、さまざまな利害衝突を解決することができる。

【0051】上述の説明は、本発明の例示的な実施形態に関してなされたものである。このような説明、および添付した図面やクレームから当業者には容易に理解されるところであるが、次のクレームで定義される本発明の

精神および範疇から逸脱することなく、多様な変化、変更、および変形を実施することができる。

【図面の簡単な説明】

【図1】 図1は、本発明の好適な実施例による音声理解およびチャンネル選局システムを示す概略図である。

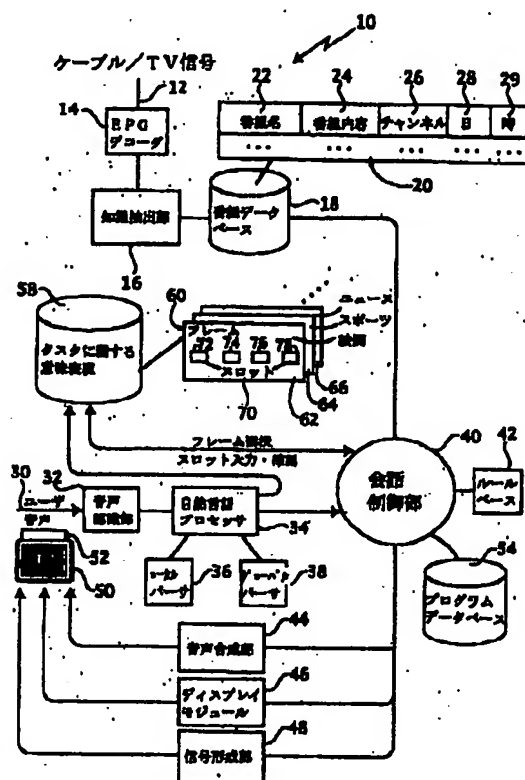
【図2】 図2は、図1で示した自然言語プロセッサおよび会話制御部による音声理解技術を示す概略図である。

【符号の説明】

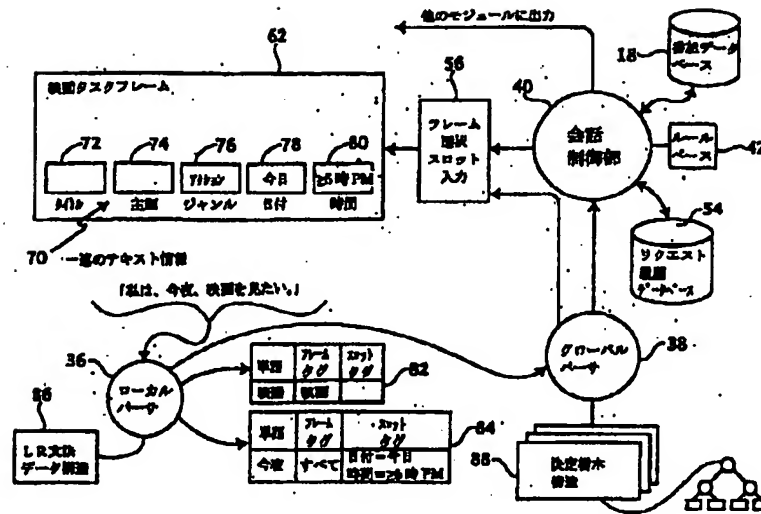
10 チャンネル選局システム、14 EPGデコー、*10

*16 知識抽出部、18 番組データベース、32 ローカルパーサ、38 グローバルパーサ、40 会話制御部、42 ルールベース、44 音声合成部、46 ディスプレイモジュール、48 信号形成部、50 テレビ受像機、52 テレビ用受信端末、58 タスクに関する意味表現集、60 タスクフレーム、70 フレームスロット、86 LR文法データ構造、88 決定樹木構造。

【図1】



【図2】



フロントページの続き

(51)Int.Cl.

H04N 7/035
7/173

識別記号

630

FI

H04N 7/08

テーマコード(参考)

A